

Qualitätssicherung mit Routinedaten im Krankenhaus:
Das Austrian Inpatient Quality Indicator System (A-IQI)
Projektabschlussbericht (Kurzbericht)

Gefördert von:

Tirol Kliniken GmbH

Tiroler Wissenschaftsförderung

Berichterstellung:

Univ.-Prof. Dr. Wolfgang Buchberger, MSc

Univ.-Ass. Marten Schmied, MA

UMIT- Private Universität für Gesundheitswissenschaften, Medizinische
Informatik und Technik

Department für Public Health, Versorgungsforschung und HTA

Institut für Qualität und Effizienz in der Medizin

November 2021

Wissenschaftliche Gesamtleitung:

Univ.-Prof. Dr. Wolfgang Buchberger, MSc.

Leiter des Instituts für Qualität und Effizienz in der Medizin
Department für Public Health, Versorgungsforschung und Health Technology Assessment
UMIT – Private Universität für Gesundheitswissenschaften, Medizinische Informatik und Technik
Eduard-Wallnöfer-Zentrum I, A-6060 Hall in Tirol
T +43 (0) 50 8648 4010, F +43 (0) 50 8648 674010
E wolfgang.buchberger@umit-tirol.at

Projektleitung:

Univ.-Ass. Marten Schmied, MA

Institut für Qualität und Effizienz in der Medizin
Department für Public Health, Versorgungsforschung und Health Technology Assessment
UMIT – Private Universität für Gesundheitswissenschaften, Medizinische Informatik und Technik
Eduard-Wallnöfer-Zentrum I, A-6060 Hall in Tirol
T +43 (0) 50 8648 4026, F +43 (0) 50 8648 674026
E marten.schmied@umit-tirol.at

Projektteam:

Univ.-Prof. Dr. Uwe Siebert, MPH, MSc

Leiter des Instituts für Public Health, Medical Decision Making und HTA
Department für Public Health, Versorgungsforschung und Health Technology Assessment
UMIT – Private Universität für Gesundheitswissenschaften, Medizinische Informatik und Technik

Wolfgang Brunner, Bakk., MSc

fhg – Zentrum für Gesundheitsberufe Tirol GmbH

Julia Hörtnagl, BScN

A.ö. Landeskrankenhaus Hall in Tirol, Innere Medizin 2

Oliver Kapferer, BScN

fhg – Zentrum für Gesundheitsberufe Tirol GmbH

OA Dr. med. Dieter Perkhofer

A.ö. Landeskrankenhaus- Universitätskliniken Innsbruck
Univ.-Klinik für Anästhesie und Intensivmedizin

OÄ Dr. med. Irma Kvitsaridze

A.ö. Landeskrankenhaus- Universitätskliniken Innsbruck
Univ.-Klinik für Strahlentherapie - Radioonkologie

Inhalt

Zusammenfassung	4
Ziele der Studie:	4
Methoden:	4
Ergebnisse	4
Schlussfolgerungen	5
Wissenschaftlicher Hintergrund	6
Ziele der Studie	8
Forschungsfragen und Methoden.....	9
Eingeschlossene PatientInnen	9
Ausschlusskriterien:	9
Zielgrößen (Endpunkte)	10
Hauptzielgrößen.....	10
Nebenzielparameter	10
Methodik.....	10
Fallzahlplanung	12
Pilotstudie	12
Ergebnisse	13
Indikatorenbezogene Auswertung.....	17
Analyse der Patientenschäden.....	18
Diskussion	25
Limitationen	27
Schlussfolgerungen	28
Referenzen	30
Anhang	32

Zusammenfassung

Ziele der Studie:

Bei dem geförderten Forschungsprojekt handelt es sich um eine Validierungsstudie des Austrian Inpatient Quality Indicator (A-IQI) Systems. Diese österreichweit eingeführten Qualitätsindikatoren auf der Grundlage von Routinedaten (LKF-Daten) sollten dahingehend untersucht werden, ob sie mit ausreichender Sicherheit das Vorliegen von unerwünschten Ereignissen (UE) im Zusammenhang mit der stationären Krankenhausbehandlung vorhersagen.

Methoden:

Im Rahmen der Studie wurden 18 im Jahr 2019 am a.ö. Landeskrankenhaus Innsbruck – Universitätskliniken auffällige A-IQI Indikatoren mit Relevanz für die Patientensicherheit untersucht. Die zugehörigen 253 Behandlungsfälle sowie 260 zufällig ausgewählte Kontrollfälle mit gleicher Falldefinition, aber unauffälligem Indikator wurden retrospektiv mit einer standardisierten zweistufigen Analyse der Krankenschichten mit dem Global Trigger Tool (GTT) als Goldstandard untersucht.

Vor Start des Projekts wurde ein Datenschutz- und Verarbeitungsvertrag mit dem Krankenanstaltenträger Tirol Kliniken abgeschlossen sowie eine Freigabe durch die Ethikkommission der Medizinischen Universität Innsbruck eingeholt. Nach Etablierung der Reviewmethode mit dem GTT an der UMIT und Schulung der jeweils aus zwei diplomierten Gesundheits- und Krankenpflegepersonen (DGKP) und einem Arzt/einer Ärztin bestehenden Reviewteams erfolgten die Reviews von Februar bis September 2021.

Die detektierten unerwünschten Ereignisse wurden für die Fall- und Kontrollgruppe als absolute Häufigkeit, absolute Häufigkeit von für das Outcome kausalen unerwünschten Ereignissen, relative Häufigkeit/1000 Behandlungstage bzw. 100 stationäre Aufnahmen und Anteil der Behandlungsfälle mit mindestens einem unerwünschten Ereignis berechnet sowie in Form von Kreuztabellen für jeden Indikator mit Angabe von Sensitivität, Spezifität und positivem sowie negativem prädiktivem Wert dargestellt. Weiters wurden Gruppenvergleiche zwischen Indikator-positiven und Indikator-negativen Fällen mit Signifikanztests (χ^2 Test und bzw. exakter Test nach Fisher, Korrelation nach Spearman) durchgeführt. Darüber hinaus wurden die detektierten unerwünschten Ereignisse kategorisiert und entsprechend den Kategorien des National Coordination Council for Medication Error Reporting and Prevention Index (NCC MERP) klassifiziert.

Ergebnisse

Ein Fall wurde wegen unvollständiger Dokumentation ausgeschlossen, 91 Indikator-positive Fälle traten bei mehr als einem Indikator auf, sodass schließlich 421 Behandlungsfälle (240 Männer und 181 Frauen zwischen 21 und 98 Jahren, Mittel: 68.6 Jahre) in die Auswertung einbezogen wurden. Durch die Chart-

Reviews wurden 161 unerwünschte Ereignisse in 86 Behandlungsfällen detektiert, das entspricht einer relativen Häufigkeit von 32,9 UE/1000 Behandlungstage bzw. 38,24 UE/100 stationäre Aufnahmen. Durchschnittlich wurden in jedem positiven Fall 1,87 unerwünschte Ereignisse gefunden, 81 (50,31%) wurden als kausal für das Indikatoren-Outcome eingeschätzt. 90 unerwünschte Ereignisse (55,9%) wurden als klinisch relevant kategorisiert (Kategorien E bis G, temporärer oder permanenter Patientenschaden), 37 (23%) als kritisch (Kategorie H, lebensrettende Intervention erforderlich) und 31 (19,3%) als zum letalen Ausgang beitragend (Kategorie I).

Über alle 18 Indikatoren wurde bei den Indikator-negativen Kontrollfällen in 10,38% und bei den Indikator-positiven Fällen in 36,64% mindestens ein unerwünschtes Ereignis detektiert ($p < 0,001$). Für die konservativen Indikatoren betrug das Verhältnis 15,79 % (Indikator negativ) vs. 32,77% (Indikator-positiv) ($p = 0,002$), für die chirurgischen Indikatoren 6,16% (Indikator-negativ) vs. 47,62% (Indikator-positiv) ($p < 0,001$). Über alle Indikatoren betrug die Sensitivität in Bezug auf die Vorhersage eines unerwünschten Ereignisses 68,08% bei einer Spezifität von 69,55%, einem positiven prädiktiven Wert von 36,65% und einem negativen prädiktiven Wert von 89,62%, wobei sich zwischen den einzelnen Indikatoren erhebliche Unterschiede zeigten.

Schlussfolgerungen

Behandlungsfälle mit auffälligem A-IQI Indikator zeigten in der durchgeführten explorativen Studie eine signifikant höhere Detektionsrate von behandlungsbedingten Schäden (unerwünschten Ereignissen) als Zufallsstichproben ohne auffälligen Indikator. Die überwiegende Mehrzahl der detektierten Schäden war von mittlerer bis hoher klinischer Relevanz.

Diese Ergebnisse lassen den Schluss zu, dass – jedenfalls für die untersuchten Indikatoren – eine Vorselektion anhand auffälliger A-IQI Indikatoren für die strukturierte Analyse von Krankengeschichten zur Qualitätssicherung in der stationären Behandlung dem konventionellen Vorgehen mit Untersuchung von zufällig ausgewählten Behandlungsfällen überlegen ist.

Wissenschaftlicher Hintergrund

Seit den beiden Reports des Institute of Medicine (IOM) „To Err is Human: Building a Safer Health System“ von 1991 und „Crossing the Quality Chasm: A New Health System for the 21st Century“ von 2001 ist Patientensicherheit zu einem prioritären Thema für Gesundheitssysteme und Gesundheitseinrichtungen geworden. Epidemiologische Studien in zahlreichen Ländern bestätigten, dass behandlungsbedingte unerwünschte Ereignisse häufig sind und schwerwiegende Konsequenzen für Patientinnen und Patienten haben können (Zegers 2009, Soop 2009, Rafter 2017, Sari 207, Classen 2013, Landrigan 2010). Nach zwei großen Metaanalysen kommt es bei 9.2% aller stationären Krankenhausbehandlungen und bei 14.4% aller stationären Behandlungen an chirurgischen Abteilungen zu mindestens einem unerwünschten Ereignis (De Vries 2008, Anderson 2013). 43.5% bzw. 37.9% dieser Ereignisse sind vermeidbar, die Mortalität beträgt in den Metaanalysen 3.6% bis 7.4%.

Unter unerwünschten Ereignissen (UE, englisch: adverse events, AE) werden generell Vorkommnisse oder Ereignisse verstanden, die möglicherweise, aber nicht zwangsläufig zu einem Schaden für die Patientinnen und Patienten führen können (Glossar Patientensicherheit, Ärztliches Zentrum für Qualität in der Medizin 2005). Das Institute of Medicine (IOM) definiert unerwünschte Ereignisse wie folgt: „An adverse event results in unintended harm to the patient by an act of commission or omission rather than by the underlying disease or condition of the patient“ (Institute of Medicine IOM 2004). In Übereinstimmung mit dieser Definition beschreibt die WHO Patientensicherheit als “reduction of risk of unnecessary harm associated with healthcare to an acceptable minimum. An acceptable minimum refers to the collective notions of given current knowledge, resources available and the context in which care was delivered weighed against the risk of non-treatment or other treatment“ (WHO 2009).

Eine Möglichkeit zur Beurteilung der Patientensicherheit ist die systematische Suche nach unerwünschten Ereignissen durch strukturierte Analyse von Krankengeschichten (Chart Review). Etablierte Chart Review-Methoden sind die Harvard Medical Practice Study (HMPS) und das vom Institute for Healthcare Improvement (IHI) entwickelte Global Trigger Tool (GTT) (Brennan et al. 1991; Griffin & Resar 2009). Chart Reviews gelten als sensitive Methoden in der Detektion von behandlungsbedingten unerwünschten Ereignissen im Krankenhaus (Health Quality & Safety Commission New Zealand 2012). Die Nachteile liegen im verhältnismäßig hohen Aufwand, der einem breiten Einsatz im klinischen Qualitäts- und Risikomanagement entgegensteht (Murff, Patel, Hripcsak & Bates 2003).

Eine andere Möglichkeit ist die Anwendung von Patientensicherheits-Indikatoren. Indikatoren sind quantitative Maße, die zur Überwachung und Bewertung der Qualität von Leitungs-, Management-, klinischen und Unterstützungsfunktionen, die sich auf die Patientenergebnisse auswirken, verwendet werden können. Ein Indikator ist jedoch kein direktes Maß für die Qualität (Joint Commission for the

Accreditation of Healthcare Indicators, JCAHO 1991). Patientensicherheits-Indikatoren (PSI) sind Qualitätsindikatoren mit einem speziellen Fokus auf Patientensicherheit. Der grundsätzliche Nachteil aller Indikatoren besteht darin, dass sie zwar auffällige Verteilungen von sicherheitsrelevanten Outcomes (z. B. Mortalität, Verlegung auf die Intensivstation, Re-Operationen) aufdecken, nicht jedoch die allenfalls zugrundeliegenden unerwünschten Ereignisse detektieren oder beschreiben können. Da jedoch nur aus der Analyse von konkreten Schadensereignissen Maßnahmen zur Prävention abgeleitet werden können, ist ihr praktischer Wert für die Qualitätssicherung begrenzt.

Ein Konzept für die Qualitätsmessung anhand von Routinedaten aus Krankenhäusern sind die durch die American Agency for Healthcare Research and Quality (AHRQ) entwickelten Inpatient Quality Indicators (IQI), welche explizit zur Messung der Versorgungsqualität entwickelt wurden und daher auch zur Identifizierung möglicher Verbesserungspotentiale geeignet sind (Farquhar, 2008). Nachdem das System an die österreichischen Bedürfnisse angepasst wurde, wurden die auf LKF-Daten beruhenden Austrian Inpatient Quality Indicators (A-IQI) im Jahr 2013 als bundesweit einheitliches Qualitätsmessverfahren für den stationären Bereich aller österreichischen Fondskrankenanstalten eingeführt. Ähnliche Indikatorensets wurden für Deutschland (D-IQI) und die Schweiz (CH-IQI) entwickelt, wodurch Ländervergleiche möglich sind.

Das A-IQI Indikatorenset in der aktuellen Version 2019 enthält 380 Indikatoren, darunter 145 Indikatoren mit Zielbereich und 17 Sentinel-Indikatoren (BMSGPK 2019). Für die Kennzahlen mit Zielbereich werden Bundesdurchschnitte oder Erwartungswerte (fallbezogen, risikoadjustiert nach Alter und Geschlecht) herangezogen. Liegt der Wert der jeweiligen Krankenanstalt außerhalb des 95%-Konfidenzintervalls, gilt dieser als signifikant auffällig („rote Ampel“). Sentinel-Indikatoren sind Patientensicherheits-Indikatoren, bei denen bereits ein einzelner Todesfall zu einer signifikanten Auffälligkeit führt.

Die Vorteile von Qualitätsindikatoren auf Basis von Routinedaten liegen im geringen Erhebungsaufwand, da diese Daten zu anderen Zwecken ohnehin erhoben werden müssen. Darüber hinaus weisen sie einen hohen Grad von Vollständigkeit auf, da sie in DRG-Systemen Grundlage der Leistungsabrechnung sind. Aus diesem Grund können auch seltene Ereignisse (sofern abrechnungsrelevant) verlässlich erfasst werden. Andererseits können zur Qualitätssicherung verwendete Abrechnungsdaten durch eine erlösoptimierte Kodierpraxis auch verfälscht werden.

Studien mit Qualitätsindikatoren auf der Grundlage von administrativen Daten haben im Vergleich zu Chart Review-Methoden eine deutlich niedrigere Sensitivität in der Detektion von unerwünschten Ereignissen ergeben (Classen 2011, Maass 2015). Allerdings wurden in diesen Studien globale Sicherheitsindikatoren wie die AHRQ Patient Safety Indicators verwendet. Indikationsspezifische Indikatoren wie die A-IQI

Indikatoren wurden hinsichtlich ihrer Sensitivität und Spezifität in der Detektion von unerwünschten Ereignissen bisher nicht evaluiert.

Ziele der Studie

Das Ziel dieser explorativen Studie war die Klärung der Forschungsfrage, ob bestimmte Indikatoren des A-IQI Systems das Vorhandensein von behandlungsbedingten unerwünschten Ereignissen (adverse events) mit ausreichender Sensitivität und Spezifität vorhersagen, damit sie als Indikatoren für Patientensicherheit verwendet werden können. Als Goldstandard sollte eine systematische Analyse von Patientendokumentationen mit dem Global Trigger Tool (GTT) dienen. Die Auswahl der Indikatoren beschränkte sich auf jene, die sich auf sicherheitsrelevante Outcomes beziehen. Dies gilt einerseits für Indikatoren mit einem definierten Zielbereich (z.B. Mortalitätsrate für eine bestimmte Hauptdiagnose oder medizinische Behandlung) und andererseits für sogenannte Sentinel-Indikatoren, bei denen schon ein einziger Todesfall zu einem auffälligen Indikator führt.

Da die A-IQI Indikatoren im Gegensatz zu globalen Qualitätsindikatoren nur jeweils für eine bestimmte Diagnose- oder Leistungsgruppe gelten, musste die Analyse für jeden untersuchten Indikator getrennt erfolgen. Daher musste sich die Untersuchung auf auffällige Indikatoren („rote Ampeln“) beschränken.

In der Studie sollte die relative Häufigkeit von durch das Global Trigger Tool identifizierten unerwünschten Ereignissen bei stationären Behandlungsfällen mit auffälligem A-IQI Indikator mit der relativen Häufigkeit von unerwünschten Ereignissen bei stationären Behandlungsfällen mit gleicher Falldefinition (Hauptdiagnose oder medizinische Einzelleistung, eventuell beschränkt auf bestimmte Altersgruppen oder Zusatzdiagnosen), jedoch ohne auffälligen Indikator verglichen werden.

Forschungsfragen und Methoden

Folgende Forschungsfragen sollten beantwortet werden:

1. Wie groß ist die Anzahl detektierter behandlungsbedingter unerwünschter Ereignisse bei Behandlungsfällen mit auffälligem Indikator im Vergleich zu einer Kontrollgruppe von Behandlungsfällen mit derselben Falldefinition, jedoch ohne auffälligen Indikator, und ist dieser Unterschied in der jeweiligen Gruppe statistisch signifikant?
2. Unterscheiden sich die detektierten unerwünschten Ereignisse in der Gruppe der Behandlungsfälle mit auffälligem Indikator in Bezug auf Art und Schweregrad von jenen in der Kontrollgruppe ohne auffälligen Indikator?
3. Wie hoch ist die Anzahl der für den Outcome kausalen unerwünschten Ereignisse bei Behandlungsfällen mit auffälligem Indikator?
4. Wie hoch sind Sensitivität, Spezifität, positiver und negativer prädiktiver Wert der untersuchten A-IQI Indikatoren in der Detektion von unerwünschten Ereignissen mit dem Global Trigger Tool (GTT) als Goldstandard?

Eingeschlossene PatientInnen

In die Studie eingeschlossen wurden Patientinnen und Patienten, die sich im Jahr 2019 einer stationären Behandlung am a. ö. Landeskrankenhaus – Universitätskliniken Innsbruck unterzogen und deren Behandlungsfall der Definition eines auffälligen Indikators („rote Ampel“) entsprach (Fallgruppe) oder deren Behandlungsfall dieselben Einschlusskriterien, jedoch keinen auffälligen Indikator aufwies (Kontrollgruppe).

Beispiel:

Fallgruppe: Patientinnen und Patienten im Alter zwischen 65 und 84 Jahren mit der Hauptdiagnose Herzinfarkt, die im Laufe des Krankenhausaufenthaltes verstarben (Indikator 01.13 Hauptdiagnose Herzinfarkt, Altersgruppe 65-84, Anteil Todesfälle- Fallgruppe).

Kontrollgruppe: Patientinnen und Patienten im Alter zwischen 65 und 84 Jahren, mit der Hauptdiagnose Herzinfarkt, die im Laufe des Krankenhausaufenthaltes nicht verstarben (Indikator 01.13 Hauptdiagnose Herzinfarkt, Altersgruppe 65-84, Anteil Todesfälle- Kontrollgruppe).

Ausschlusskriterien:

Gemäß den Anwendungskriterien des Global Trigger Tool (Griffin 2009, Herold 2013) wurden folgende Patientinnen und Patienten nicht in die Studie eingeschlossen:

- A) Patientinnen und Patienten, die zur Zeit der stationären Aufnahme jünger als 18 Jahre waren;

- B) die für weniger als 24 Stunden stationär aufgenommen waren und nicht im Krankenhaus verstarben;
- C) die (zum Stichtag) vor weniger als 30 Tagen entlassen wurden oder;
- D) deren Hauptdiagnose eine psychiatrische war.

Darüber hinaus wurden für jeden Indikator diejenigen PatientInnen ausgeschlossen, die nicht der Falldefinition (z.B. definierte Altersgruppe) entsprechen.

Zielgrößen (Endpunkte)

Im Rahmen dieser Studie wurden verschiedene Zielgrößen betrachtet, die Aussagen zur Validität der untersuchten A-IQI als Indikatoren für Patientensicherheit zulassen.

Hauptzielgrößen

1. Sensitivität, Spezifität, positiver und negativer prädiktiver Wert bezüglich der Vorhersage von unerwünschten Ereignissen für die einzelnen Indikatoren sowie für alle Indikatoren zusammen;
2. Relative Häufigkeit von unerwünschten Ereignissen für alle Indikatoren entsprechend der Systematik des Global Trigger Tools wird die relative Häufigkeit berechnet als:
 - prozentualer Anteil der Behandlungsfälle mit mindestens einem unerwünschten Ereignis;
 - Anzahl der kumulativen unerwünschten Ereignisse/1000 Behandlungstage;
 - Anzahl der kumulativen unerwünschten Ereignisse/100 stationäre Aufnahmen.

Nebenzielparameter

Eine Reihe von Nebenzielparametern wurden bei der retrospektiven Chart Review erhoben und konnten zur Beurteilung der eingesetzten Methodik sowie für die Identifikation von Risiken bzgl. der Patientensicherheit herangezogen werden. Als Nebenzielparameter erhoben werden sollten:

1. Alter und Geschlecht von Patientinnen und Patienten mit/ohne unerwünschtem Ereignis für alle Indikatoren zusammen;
2. Anzahl der als kausal für den Outcome eingeschätzten unerwünschte Ereignisse für jeden Indikator und für alle Indikatoren zusammen;
3. Klassifizierung der identifizierten unerwünschten Ereignisse entsprechend der Kategorien des National Coordination Council for Medication Error Reporting and Prevention Index (NCC MERP).

Methodik

Im Rahmen der Studie wurden retrospektive Fallanalysen (Chart Reviews) unter Verwendung der elektronischen Krankenakten, Ambulanzakten und OP-Berichte der Tirol Kliniken zur Detektion unerwünschter Ereignisse durchgeführt. Die Chart Reviews wurden entsprechend der Vorgehensweise des

durch das Institute for Healthcare Improvement (IHI) entwickelten Global Trigger Tools (GTT) durchgeführt (Griffin 2009, Herold 2013).

Die Auswahl der Behandlungsfälle erfolgte anhand der im Jahr 2019 am a. ö. Landeskrankenhaus – Universitätskliniken Innsbruck auffälligen A-IQI Indikatoren, entsprechend den beschriebenen Kriterien. Alle Behandlungsfälle mit auffälligem Indikator wurden mit dem GTT analysiert. Aus der Kontrollgruppe wurden zufällige Stichproben für die strukturierte Analyse der Krankengeschichten mit dem GTT gezogen.

Die Krankengeschichten wurden zunächst von zwei Reviewern (diplomierte Gesundheits- und Krankenpflegepersonen) anhand eines standardisierten Auswerteschemas nach Hinweisen auf das Vorliegen eines unerwünschten Ereignisses („Trigger“) gescreent. Zunächst führten die beiden Reviewer voneinander unabhängig die Analyse durch und dokumentieren ihre Ergebnisse. Für die Durchsicht einer Krankengeschichte sollten dabei maximal 20 Minuten verwendet werden (Griffin 2009, Herold 2013).

Die Dokumentation der Ergebnisse erfolgte auf einem Datenblatt (Case Report Form, siehe Anhang A) in pseudonymisierter Form. Für die Dokumentation der Chart-Reviews wurde die deutschsprachige Version der GTT-Vorlage des IHI genutzt (Institute of Healthcare Improvement IHI, Cambridge MA, USA 2009). Die Reviewer dokumentierten ihre Ergebnisse zunächst getrennt und führten dann im Rahmen einer Konsensuskonferenz ihre Ergebnisse zusammen und entschieden gemeinsam über das Ergebnis, welches dann den ärztlichen Supervisor vorgelegt wurde.

Der Supervisor (Arzt mit entsprechender klinischer Erfahrung) entschied anhand der Dokumentation der Reviewer über Vorliegen, Art und Schweregrad eines unerwünschten Ereignisses. Im Bedarfsfall nahm der Supervisor ebenfalls Einsicht in die Patientendokumentation.

Ob die einzelnen Fälle aus der Interventions- oder Kontrollgruppe stammen, war ausschließlich aus dem ursprünglichen Zuteilungsblatt ersichtlich, welches von der Projektleitung zunächst angelegt, während der eigentlichen Chart-Reviews unter Verschluss gehalten und erst für die abschließende Datenauswertung konsultiert wurde. Eine vollständige Verblindung von Reviewer und Supervisor war dennoch nicht möglich da im Rahmen des Chart Reviews in den meisten Fällen festgestellt werden konnte, ob die betroffenen PatientInnen verstorben waren (Fallgruppe) oder überlebt hatten (Kontrollgruppe).

Alle in der Studie erhobenen Daten aus den Chart-Reviews wurden zunächst in Excel Dateien gesammelt und für die statistische Auswertung in IBM SPSS Statistics 27 übertragen. Die für die Indikatoren erstellten Kreuztabellen (4-Felder-Tafeln) wurden mittels eines Chi²-Tests auf signifikante Unterschiede in der Verteilung der unerwünschten Ereignisse zwischen Verdachts- und Kontrollgruppe untersucht. Für die Signifikanztest der Indikatoren, deren Stichprobenumfang ≤ 5 betrug, wurde der exakte Test nach Fisher angewendet, da dieser auch bei einer geringen Anzahl von Beobachtungen zuverlässige Resultate liefert.

Zur Beurteilung der Signifikanz (p -Wert) wurde eine Irrtumswahrscheinlichkeit $< 5\%$ zu Grunde gelegt. Schließlich wurde für alle Indikatoren und die Indikatorengruppen die Rangkorrelationskoeffizient nach Spearman berechnet, um eine Aussage über die Korrelation zwischen dem Vorliegen eines auffälligen Indikators und dem Auftreten von mindestens einem unerwünschten Ereignis zu treffen.

Fallzahlplanung

Für das Jahr 2019 wurden am a. ö. Landeskrankenhaus – Universitätskliniken Innsbruck im Rahmen der routinemäßigen Qualitätssicherung 19 auffällige A-IQI-Indikatoren festgestellt. Davon bezogen sich 18 Indikatoren (12 Indikatoren mit Erwartungswert bzw. österreichweitem Durchschnittswert als Zielbereich und 6 Sentinel-Indikatoren) auf sicherheitsrelevante Outcomes. Aus insgesamt 4928 Behandlungsfällen mit der jeweiligen Falldefinition wurden von den Tirol Kliniken mit Hilfe des Programms QDoc 253 Behandlungsfälle identifiziert, für die einer der betreffenden Indikatoren positiv war.

Für den Vergleich der Proportion detektierter unerwünschter Ereignisse bei Behandlungsfällen mit einem auffälligen A-IQI Indikator und der Proportion detektierter unerwünschter Ereignisse in der Kontrollgruppe ohne auffälligen Indikator wurde eine Fallzahlschätzung durchgeführt. Die Fallzahlschätzung wurde für jeden Indikator getrennt vorgenommen. Es sollten alle positiven Fälle in die Auswertung eingehen. Ziel der Fallzahlschätzung war somit die Ermittlung der Anzahl der Kontrollen, die den Fällen gegenübergestellt werden sollten und die hinsichtlich detektierter unerwünschter Ereignisse geprüft werden sollten.

Für die Fallzahlschätzung wurde ein einseitiger χ^2 -Test für den Vergleich von zwei unabhängigen Proportionen verwendet. Als Signifikanzlevel des Tests wurde 0,05 verwendet und als Power 80%.

Pilotstudie

Zur Abschätzung der zu erwartenden Prävalenz von unerwünschten Ereignissen bei Behandlungsfällen mit auffälligem A-IQI Indikator wurde eine Pilotstudie von 20 Indikator-positiven stationären Behandlungsfällen an den Tirol Kliniken aus dem Jahr 2017 durchgeführt. Diese ergab eine Gesamtproportion detektierter unerwünschter Ereignisse von 0,60 (95%-Konfidenzintervall 0,39-0,81). Für die Fallzahlschätzung wurde eine Proportion in der Fallgruppe von 0,50 angenommen. Basierend auf der Literatur (De Vries 2008) wurde in der Kontrollgruppe eine Proportion detektierter unerwünschter Ereignisse von 0,092 angenommen.

Aus Effizienzgründen wurde ein maximales Verhältnis von Kontrollen zu Fällen von 4 :1 festgelegt. Zudem wurde minimales Verhältnis Kontrollen zu Fällen von 1:2 festgelegt, um auch in den Indikatorgruppen mit sehr vielen Fällen eine ausreichende Zahl an Kontrollen zu erhalten. Daraus ergab sich eine Anzahl von 253 Fällen und 260 Kontrollen (insgesamt: 513) für die Chart Reviews. Eine detaillierte Auflistung der in die Studie eingeschlossenen Fälle sowie Kontrollen pro Indikator findet sich im Anhang B dieses Berichtes.

Ergebnisse

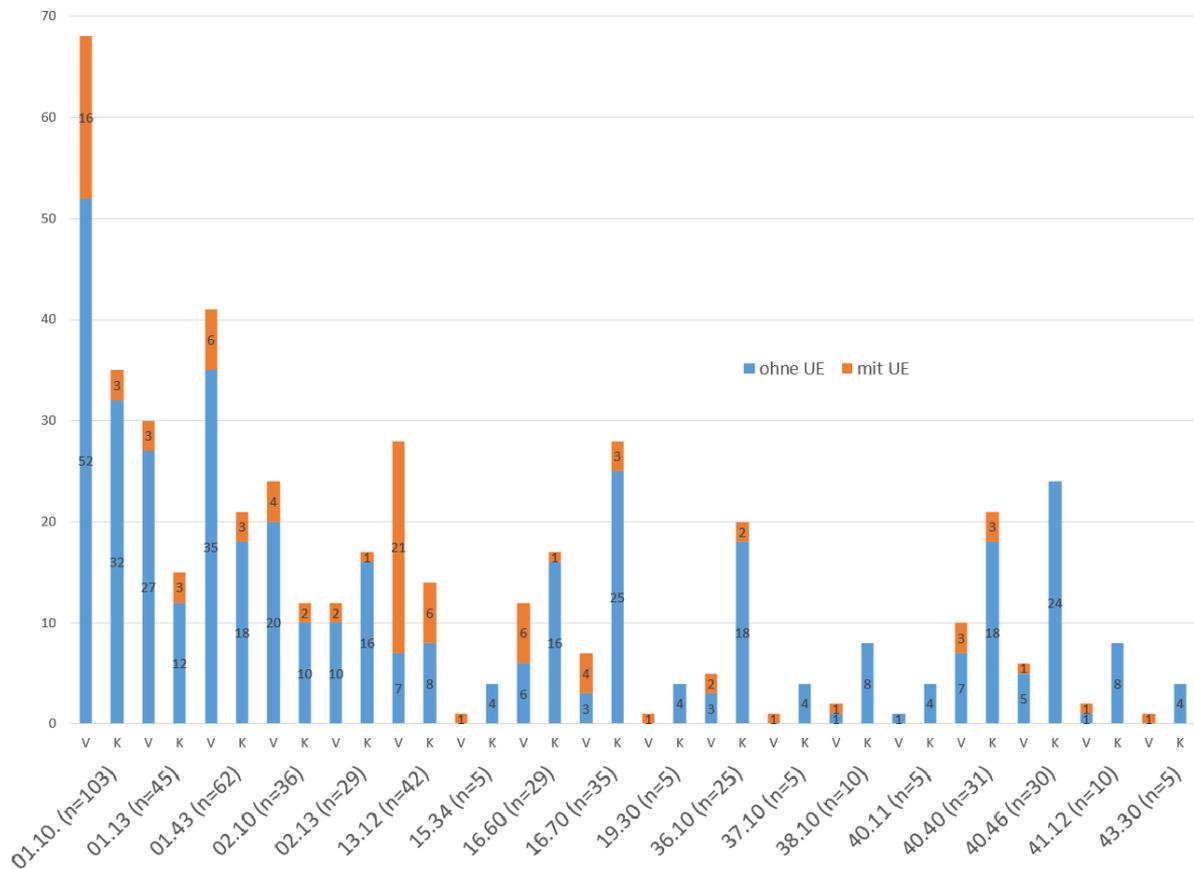
Um die Bearbeitung der Chart Reviews in einem angemessenen zeitlichen Rahmen durchführen zu können wurden die Indikatoren in zwei Gruppe geteilt. Die Indikatoren Hauptdiagnose Herzinfarkt (01.10, 01.13 & 01.43), Hauptdiagnose Herzinsuffizienz (02.10 & 02.13) sowie Pneumonie, im stationären Aufenthalt erworben (13.12) bildeten dabei die erste Gruppe der *konservativen Indikatoren*. Die zweite Gruppe bildeten die verbleibenden zwölf *chirurgischen Indikatoren* (15.34, 16.60, 16.70, 19.30, 36.10, 37.10, 38.10, 40.40, 40.46, 41.12 & 43.30).

Entsprechend der Fallzahlplanung wurden in die Auswertung des Projektes 513 Chart Reviews, 253 Fälle und 260 Kontrollen, aufgenommen. Während der Chart-Reviews stellte sich heraus, dass die Datenlänge zu einem Fall aus Indikator 01.10 nicht ausreichend für eine abschließende Beurteilung war, weshalb dieser Fall entfernt wurde. In den verbliebenen 512 Chart Reviews fanden sich weiters 91 Falldubletten, d.h. Fälle, die sich auf mehr als einen Indikator bezogen. Da keine zusätzlichen, zu einem roten Indikator beitragenden Behandlungsfälle aus dem Jahr 2019 zur Verfügung standen, wurden diese Duplikate aus der statistischen Analyse über alle Indikatoren und die Indikatorgruppen entfernt, jedoch in der Analyse der einzelnen Indikatoren belassen. Insgesamt wurden 421 Chart-Reviews (512 minus 91 Duplikate) durchgeführt und in der statistischen Analyse über alle Indikatoren berücksichtigt. In die Auswertung der konservativen Indikatorengruppe konnten 233 Chart-Reviews (317 minus 84 Duplikate), für die chirurgische Indikatorengruppe 188 Chat-Reviews (195 minus 7 Duplikate) aufgenommen werden. Für die Auswertung auf Ebene der einzelnen Indikatoren verblieben 512 Chart-Reviews. Anhang B dieses Berichtes zeigt eine genaue Aufschlüsselung der entfernten Falldubletten und der in der Analyse verbliebenen Chart-Reviews per Indikator.

In den 421 eingeschlossenen Falldokumentationen fanden sich 57% (n=240) Männer und 43% (n=181) Frauen. Das durchschnittliche Alter betrug 68,57 (SD: 17,27, Spannweite: 21- 98 Jahre). Der stationäre Aufenthalt der Patientinnen und Patienten dauerte zwischen 0 und 119 Tage, im Durchschnitt 11,62 (SD:17,44) Behandlungstage. Für die konservativen Indikatoren betrug der Aufenthalt im Mittel 10,72 Tage (SD:18,16), für die chirurgischen Indikatoren 12,74 (SD:16,5) Tage.

Durch die Chart-Reviews zwischen März und Juli sowie die abschließenden Konsenskonferenzen im Herbst 2021 wurden 161 unerwünschte Ereignisse in 86 Behandlungsfällen identifiziert. Abbildung 1 zeigt die Verteilung der in die statistische eingeschlossenen 512 Behandlungsfälle auf die Indikatoren entsprechend ihrer Beurteilung durch die Chart-Review anhand des GTT.

Abb.1: Verteilung der Verdachtsfälle (V) mit positivem Indikator und Kontrollfälle (K) mit negativem Indikator auf alle Behandlungsfälle (n=512) nach Indikatoren und Vorliegen von mindestens einem unerwünschten Ereignis (UE)



Die folgende Tabelle 1 gibt eine Übersicht der im Projekt festgestellten testtheoretischen Gütekriterien (Sensitivität, Spezifität, positiver und negativer prädiktiver Wert), den prozentualen Anteil von Behandlungsfällen gesamt mit mindestens einem unerwünschten Ereignis, den prozentualen Anteil von Behandlungsfällen mit mindestens einem unerwünschten Ereignis unter den Indikator-positiven (Verdachtsfälle) und Indikator-negativen Fällen (Kontrollfälle), die Rate von unerwünschten Ereignissen pro 1000 Behandlungstage und pro 100 stationäre Aufnahmen. Weiters werden die Ergebnisse der Signifikanztests für den Vergleich von Indikator-positiven Verdachtsfällen und Indikator-negativen Kontrollfällen angegeben.

Tab. 1: Indikatorenbezogene Auswertung

Indikator	Sensitivität	Spezifität	PPV	NPV	Behandlungsfälle mit mindestens einem unerwünschten Ereignis	Verdachtsfälle mit mindestens einem unerwünschten Ereignis	Kontrollfälle mit mindestens einem unerwünschten Ereignis	AE/1000 Behandlungstage	AE/100 Aufnahmen	Korrelation nach Spearman (Näherungsweise Signifikanz)
Indikatoren insgesamt (n=421)	68,60%	69,55%	36,65%	89,62%	20,43% (86)	36,64% (59)	10,38% (27)	32,9	38,24	0,317 (0,000**)
Indikatoren konservativ (n=233)	68,42%	54,55%	32,77%	84,21%	24,46% (57)	32,77% (39)	15,79% (18)	47,26	50,64	0,197 (0,002**)
01.10. Hauptdiagnose Herzinfarkt (Alter > 19), Anteil Todesfälle (n=103)	84,21%	38,10%	23,53%	91,43%	18,45% (19)	23,53% (16)	8,57% (3)	72,02	25,24	0,183 (0,065)
01.13 Hauptdiagnose Herzinfarkt, Altersgruppe 65-84, Anteil Todesfälle (n=45)	50%	30,77%	10%	80%	13,34% (6)	10% (3)	20% (3)	45,45	22,23	-0,139 (0,364)
01.43 Hauptdiagnose transmuraler Herzinfarkt (Alter > 19), Anteil Todesfälle (n=62)	66,67%	33,96%	14,63%	85,71%	14,52% (9)	14,63% (6)	14,28% (3)	54,6	30,64	0,005 (0,971)
02.10 Hauptdiagnose Herzinsuffizienz (Alter > 19), Anteil Todesfälle (n=36)	66,67%	33,33%	16,67%	83,33%	16,67% (6)	16,67% (4)	16,67% (2)	27,17	27,78	0,000 (1)
02.13 Hauptdiagnose Herzinsuffizienz, Altersgruppe 65-84, Anteil Todesfälle (n=29)	66,67%	61,54%	16,67%	94,12%	10,34% (3)	16,67% (2)	5,88% (1)	27,27	20,68	0,174 (0,366)
13.12 Pneumonie im stat. Aufenthalt erworben (hospital-acquired pneumonia), Anteil Todesfälle (n=42) ***	77,78%	53,33%	75%	57,14%	64,28% (27)	75% (21)	42,85% (6)	50,63	171,43	0,357 (0,020*)
Indikatoren chirurgisch (n=188)	68,97%	86,16%	47,62%	93,84%	15,42% (29)	47,62% (20)	6,16% (9)	17,95	22,87	0,478 (0,000**)
15.34 Teilresektionen der Lunge (ohne Lobektomie) bei Bronchialkarzinom, Anteil Todesfälle (n=5)	100%	100%	100%	100%	20% (1)	100% (1)	0% (0)	37,04	40	1 (0,200 ▲)
16.60 Laparoskopische Cholezystektomien ohne Tumorfälle, Anteil Umsteiger (n=29)	85,71%	72,73%	50%	94,12%	24,14%(7)	50% (6)	5,88% (1)	26,24	31,03	0,508 (0,005**)
16.70 Laparoskopische Cholezystektomien ohne Tumorfälle, Anteil ERCP nach OP-Tag (n=35)	57,14%	80,65%	40%	89,29%	20% (7)	57,14% (4)	10,71% (3)	36,46	20	0,464 (0,005**)
19.30 Adipositas-Chirurgie, Anteil Todesfälle (n=5)	100%	100%	100%	100%	20% (1)	100% (1)	0% (0)	35,71	40	1 (0,200 ▲)
36.10 Hüft-Endoprothesen-Revisionen insgesamt, Anteil Todesfälle (n=25)	50%	85,71%	40%	90%	16% (4)	40% (2)	10% (2)	13,06	36	0,327 (0,11)

37.10 Knie-Endoprothesen-Erstimpl. b. Gonarthrose u. chronischer Arthritis, Anteil Todesfälle (n=5)	100%	100%	100%	100%	20% (1)	100% (1)	0% (0)	32,26	20	1 (0,200 ▲)
38.10 Knie-Endoprothesen-Revisionen insgesamt, Anteil Todesfälle (n=10)	100%	88,89%	50%	100%	10% (1)	50% (1)	0% (0)	4,02	10	0,667 (0,035*)
40.11 Hüftgelenknahe Fraktur, Altersgruppe 20-59, Anteil Todesfälle (n=5) ****	NA	80%	NA	100%	0%	0% (0)	0% (0)	0	0	NA
40.40 Pertrochantäre Fraktur (Alter > 19), Anteil Todesfälle (n=31)	50%	72%	30%	85,71%	19,35% (6)	33,34% (3)	14,29% (3)	33,13	35,48	0,186 (0,317)
40.46 Pertrochantäre Fraktur, Altersgruppe >= 90, Anteil Todesfälle (n=30)	100%	82,76%	16,67%	100%	3,33% (1)	16,67% (1)	0% (0)	4,1	3,34	0,371 (0,043*)
41.12 Operationen an der Wirbelsäule und am RM (ohne Tumor, Trauma), Anteil Todesfälle (n=10)	100%	88,89%	50%	100%	10% (1)	50% (1)	0% (0)	14,42	30	0,667 (0,035*)
43.30 Partielle Nephrektomien bei bösartiger Neubildung der Niere, Anteil Todesfälle (n=5)	100%	100%	100%	100%	20% (1)	100% (1)	0% (0)	44,45	40	1 (0,200 ▲)

*signifikant

** hoch signifikant

▲ für Stichprobengrößen ≤5 wurde der exakte Test nach Fisher angewandt

*** im KH erworbenen Pneumonie nicht als Patientenschaden berücksichtigt

**** in der Stichprobe wurden keine Patientenschäden festgestellt

Indikatorenbezogene Auswertung

Sensitivität, Spezifität, positiver (PPW) und negativer prädiktiver Wert (NPW) der untersuchten A-IQI Indikatoren zeigten insgesamt moderate bis gute Werte. Die Ergebnisse der Chart Reviews über die achtzehn Indikatoren insgesamt (n=421) ergaben eine Sensitivität für die Vorhersage von unerwünschten Ereignissen von 68,08% bei einer Spezifität von 69,55%, einem positiven prädiktiven Wert von 36,65% und einem negativen prädiktiven Wert von 89,62%.

Bei den konservativen Indikatoren schwankte die Sensitivität zwischen 84,21% (Indikator 01.10) und 50% (Indikator 01.13) und lag damit jeweils etwas höher als die Spezifität zwischen 61,54% (Indikator 02.13) und 30,77% (Indikator 01.13). Während der PPW für den Indikator 01.13 mit 10% sehr niedrig war, wurden bei dem Indikator 13.12 75% der Verdachtsfälle positiv auf ein UE getestet. Der NPW lag zwischen 80% (Indikator 01.13) und 94,43% (Indikator 01.10) und war mit Ausnahme des Indikators 13.12 (57,14%) in durchwegs höher als der PPW. Dabei kommt dem Indikator 13.12 (Pneumonie im stationären Aufenthalt erworben) eine Sonderstellung auch deshalb zu, da per Definition jede Chart-Review mit dem GTT positiv auf ein unerwünschtes Ereignis (Trigger G11 oder I1) hätte testen müssen.

Bei den chirurgischen Indikatoren erreichten, vermutlich aufgrund der jeweils geringen Stichprobengröße von $n \leq 5$, vier Indikatoren (15.34, 19.30, 37.10 & 43.30) perfekte Ergebnisse von 100% für Sensitivität, Spezifität, NPW und PPW. Für den Indikator 40.11 konnte keine Sensitivität berechnet werden, da in diesem Sample (n=5) kein UE festgestellt wurde. Die verbleibenden vier chirurgischen Indikatoren hatten eine Sensitivität zwischen 50% (Indikator 36.10 & 40.40) und 85,71% sowie eine Spezifität zwischen 72,73% (Indikator 16.60) und 88,89% (Indikatoren 41.12 & 38.10). Eine detaillierte Darstellung der Vier-Felder-Tafeln für alle untersuchten Indikatoren findet sich im Anhang C.

Die proportionale Verteilung von Behandlungsfällen mit mindestens einem UE war, mit Ausnahme des Indikators 01.13 in der Fallgruppe stets höher als in der Kontrollgruppe. In den Kontrollfällen wurden über alle Indikatoren 10,38% (27) positiv auf mindestens ein UE getestet, welches der erwarteten Prävalenz aus den Vorstudien (De Vries 2008, Anderson 2013) entspricht. Bei den Verdachtsfällen wurde mit 36,64% (27) bei mehr als einem Drittel der Chart-Reviews durch das GTT mindestens ein UE festgestellt. Auch für die konservativen Indikatoren wurde knapp ein Drittel (32,77% (39)) der Verdachtsfälle positiv auf mindestens ein UE getestet, während es hier in der Kontrollgruppe 15,79% (39) waren. Für die chirurgischen Indikatoren waren es nur 6,16% (9) unter den Kontrollfällen, während fast die Hälfte (47,62% (20)) der Verdachtsfälle positiv mit dem GTT getestet wurde.

Um die Frage, ob der Unterschied in der Verteilung der Behandlungsfälle mit mindestens einem UE sich in Verdachts- und Kontrollgruppe signifikant unterscheidet, wurde ein Chi²-Test für zwei unverbundene Stichproben durchgeführt sowie die Korrelation nach Spearman berechnet. Hoch signifikant ($p < 0,01$) war

demnach der Unterschied in der Verteilung der AE-positiven Fälle über alle Indikatoren sowie über die beiden Indikatorengruppen (chirurgische und konservative Behandlungsfälle). Auf Ebene der einzelnen Indikatoren war das Ergebnis für den konservativen Indikator 13.12 und die chirurgischen Indikatoren 38.10, 40.46 & 41.12 signifikant ($p < 0,05$), sowie für die Indikatoren 16.60 (Laparoskopische Cholezystektomien Anteil Umsteiger) 16.70 (ERCP nach OP-Tag) hoch signifikant.

Entsprechend der Systematik des Global Trigger Tools wurde außerdem das Auftreten von UE auf 1000 Behandlungstage bzw. 100 stationäre Aufnahmen berechnet. Über alle Indikatoren wurden 161 unerwünschte Ereignisse auf 4893 Behandlungstage bzw. 421 Aufnahmen festgestellt, welches 32,9 UE auf 1000 Behandlungstage bzw. 38,24 UE auf 100 Aufnahmen entspricht. Die konservativen Indikatoren enthielten mit 47,26 UE/1000 Behandlungstage bzw. 50,64 UE/100 Aufnahmen deutlich mehr UE als die chirurgischen Indikatoren mit 17,95UE/1000 Behandlungstagen bzw. 22,87 UE/ 100 Aufnahmen.

Tabelle 2 zeigt die Zusammensetzung der Studienpopulation bezogen auf Alter und Geschlecht der Patientinnen mit und ohne unerwünschtes Ereignis:

Tab. 2: Alters- und Geschlechtsverteilung für Behandlungsfälle mit und ohne unerwünschtes Ereignis

	PatientInnen ohne unerwünschtes Ereignis (n=335)			PatientInnen mit unerwünschtem Ereignis (n=86)		
	Minimum	Maximum	Mittelwert	Minimum	Maximum	Mittelwert
Alter	21	98	67,91 (SD:17,84)	22	98	70,91 (SD:15,51)
	% (n)			% (n)		
Gültig	m	55,2% (185)		m	64% (55)	
	w	44,8% (150)		w	36% (31)	

Die Behandlungsfälle mit oder ohne UE unterschieden sich im Hinblick auf die Alters- oder Geschlechtsverteilung nicht signifikant ($p > 0,05$).

Analyse der Patientenschäden

Durch die im Projekt durchgeführten 421 Charte-Reviews wurden insgesamt 161 unerwünschte Ereignisse in 86 PatientInnen detektiert. Durchschnittlich wurden damit in jedem positiven Fall 1,87 unerwünschte Ereignisse gefunden. Von den 161 festgestellten unerwünschten Ereignissen waren 81 (50,31%) kausal für das Indikatoren-Outcome.

Abbildung 2 zeigt die Anzahl der in der Verdachts- bzw. Kontrollgruppe der jeweiligen Indikatoren gefundenen UE. Da dreißig UE in Falldoubletten enthalten waren, ist die gesamte Anzahl der Schäden in dieser Auswertung n=191. Die meisten UE (n=72) wurden in Indikator 13.12 festgestellt, sechzig davon in der

Verdachtsgruppe. In den Behandlungsfällen für Indikator 40.11 (n=5) wurde weder in der Verdachts- noch in der Kontrollgruppe ein UE gefunden. Nur bei den Indikatoren 01.34 und 36.10 wurden in der Kontrollgruppe mehr UE als in der Verdachtsgruppe gefunden. Bei sieben der achtzehn Indikatoren wurden ausschließlich in der Verdachtsgruppe UE festgestellt.

Abb.2: Anteil der Fälle mit unerwünschten Ereignissen in der Verdachtsgruppe (Indikator-positiv) und Kontrollgruppe (Indikator-negativ)

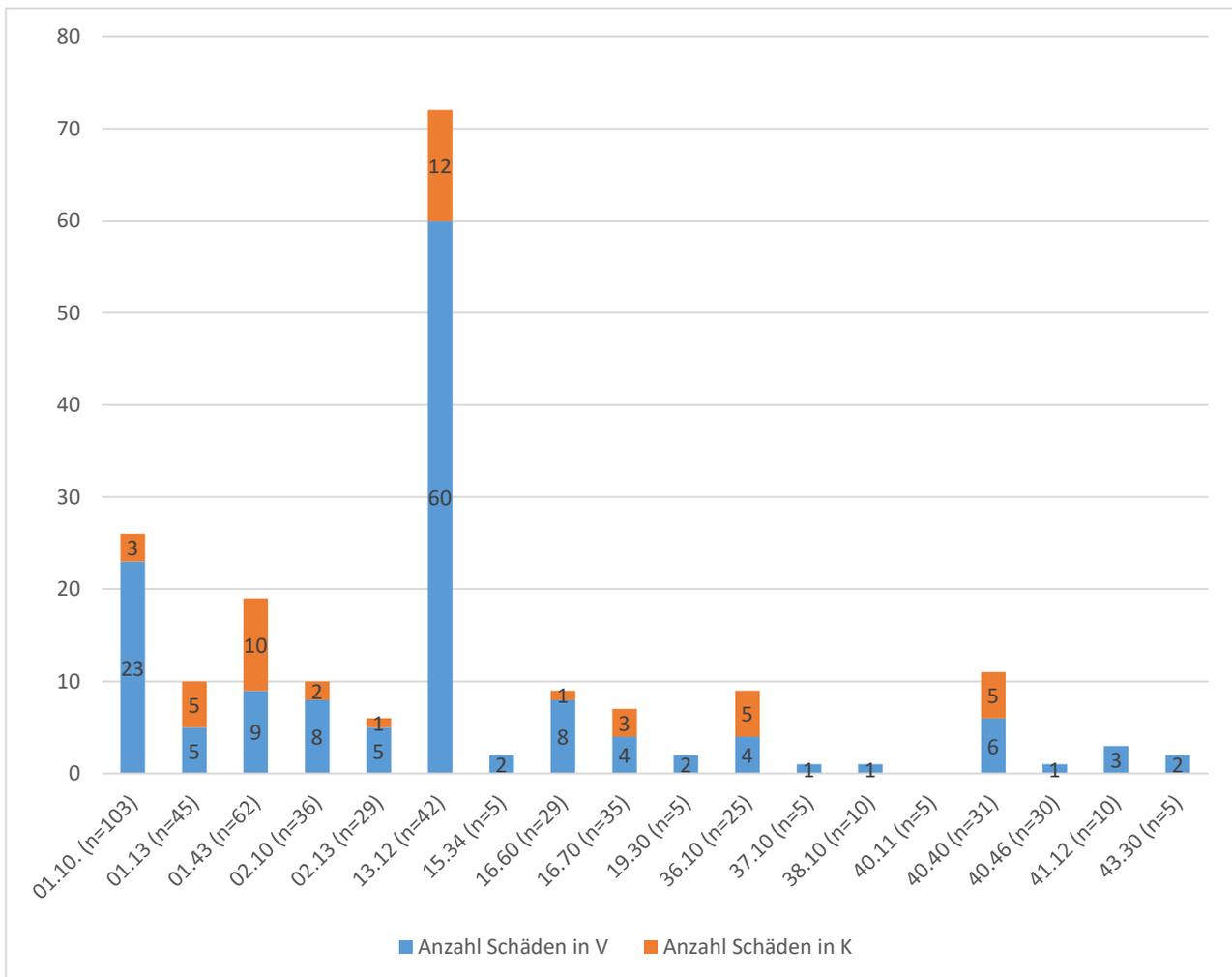


Tabelle 3 zeigt die Schadensklassifikation, entsprechend der Kategorien des National Coordination Council for Medication Error Reporting and Prevention Index (NCC MERP) und deren Auftreten über alle Chart-Reviews (n=421), die Anzahl der PatientInnen, in denen sich diese Schäden fanden und den Prozentsatz der für das Indikatoren-Outcome kausalen Patientenschäden.

Tab.3: Klassifikation der detektierten Behandlungsschäden

Schadensklassifikationen	Anzahl Schäden (%)	Anzahl Behandlungsfälle mit Schaden	Kausal für das Outcome
C -Ein Fehler, der den Patienten zwar erreicht, aber ihm keinen Schaden zugefügt hat.	0,62% (n=1)	1	0 -
D Ein Fehler, der den Patienten erreicht hat und der eine Überwachung oder Intervention erforderte, um sicherzustellen, dass der Fehler dem Patienten keinen Schaden zugefügt hat.	1,24% (n=2)	2	0 -
E Temporärer Patientenschaden, der eine Intervention erfordert.	20,50% (n=33)	32	27,27% (n=9)
F Temporärer Patientenschaden, der einen Krankenhausaufenthalt oder eine Verlängerung des Krankenhausaufenthaltes erfordert.	30,43% (n=49)	35	18,37% (n=9)
G Permanenter Patientenschaden.	4,97% (n=8)	7	37,50% (n=3)
H Patientenschaden, der eine lebenserhaltende Intervention notwendig macht.	22,98% (n=37)	22	78,38% (n=29)
I Patientenschaden, der zum Tod des Patienten beigetragen hat.	19,25% (n=31)	26	100% (n=31)
Σ	100% (n=161)		50,31% (n=81)

Besonders in den Schadensklassifikation F, H und I zeigten sich mehrere Schäden derselben Klassifikation in denselben Behandlungsfällen. So wurden 49 Schäden der Klassifikation F in 35 unterschiedlichen Behandlungsfällen festgestellt, für die Kategorie H waren es 37 Schäden in nur 22 Behandlungsfällen und die höchste Schadensklassifikation I wurde in 26 unterschiedlichen PatientInnen insgesamt 31-mal entdeckt. Mit steigendem Schweregrad der UE steigt tendenziell auch der Anteil der für das Indikatoren-Outcome kausalen UE. So waren 27,27% (9) UE der Klasse E kausal, während es in Klasse H bereits 78,38% (29) waren und in der Klasse I 100% der 31 Schäden zum Tod des Patienten beigetragen haben und somit als kausal anzusehen waren.

Des Weiteren wurden die 161 identifizierten Schäden kategorisiert und, wo anwendbar, Subkategorien gebildet. Tabelle 4 zeigt die Verteilung der unerwünschten Ereignisse nach diesen (Sub-)Kategorien:

Tab. 4: Schadenskategorien und Subkategorien

Schadenskategorie		Subkategorie	
Im Krankenhaus erworbene Infektion	32,30% (52)	Pneumonie	32,69% (17)
		Sepsis	15,38% (8)
		Wundinfekt	13,46% (7)
		Sonstige	38,46% (20)
Behandlungsfehler, nicht näher bezeichnet	19,88% (32)	Akute behandlungsbedingte Niereninsuffizienz, dialysepflichtig	9,38% (3)
		Sonstige	90,63% (29)
Sonstige chirurgische Behandlungskomplikationen	14,29% (23)	Verletzung von Organen oder Gefäßen	34,78% (8)
		Notwendige Re-Operation	34,78% (8)
		Sonstige	30,43% (7)
Pflegerfehler	11,8% (19)	Dekubitus	42,11% (8)
		Sturz	31,58% (6)
		Sonstige	26,32% (5)
Blutung/Nachblutung	10,56% (17)		
Delir	4,97% (8)		
Komplikationen im Zusammenhang mit der Medikation	3,73% (6)		
Zu frühe oder akzidentielle Extubation	2,48% (4)		
Σ	100% (161)		

Die mit 32,3% (n= 52) häufigsten unerwünschten Ereignisse ließen sich dem Bereich der im Krankenhaus erworbenen Infektionen zuordnen. Von diesen Infektionen waren 32,69% (n= 17) Pneumonien, 15,38% (n= 8) positive Blutkulturen/Sepsis und 13,46% (n= 7) Wundinfekte.

19,88% (n= 32) der unerwünschten Ereignisse waren in ihrer Ursache mehr einem übergeordneten Versagen in der Behandlungscoordination als einem bestimmten klinischen Procedere zuzuordnen und wurden daher als „Behandlungsfehler, nicht näher bezeichnet“ kategorisiert. In dieser Kategorie finden sich unter anderem Fälle von postoperativ aufgetretenem Lungenödem, ein perioperativer Myokardinfarkt, Beinvenenthrombosen sowie im Krankenhaus aufgetretene dialysepflichtige Niereninsuffizienzen.

Blutungskomplikationen nach Eingriffen umfassten 10,56% der unerwünschten Ereignisse (n= 17). Als sonstige chirurgische Behandlungskomplikationen wurden jene 14,29% (n= 23) der unerwünschten Ereignisse kategorisiert, welche durch eine intraoperative Verletzung (Perforation von Organen oder Ruptur von Gefäßen) bedingt waren oder bei denen eine zweite Operation, zum Beispiel aufgrund behandlungsbedürftiger Steinresiduen nach Cholezystektomie, notwendig war.

Pflegefehlern (11,8%, n= 19) umfassten vor allem Stürze und Decubiti. 4,97% (8) der unerwünschten Ereignisse waren durch ein im Krankenhaus erworbenes Delir, nur 3,73% (n= 6) durch medikamentöse Behandlungsfehler bedingt. Vier unerwünschte Ereignisse (2,48%) waren mit der In- bzw. Extubation der PatientInnen assoziiert, wovon in drei Fällen offenbar zu früh, einmal akzidentiell extubiert wurde. Eine vollständige und detaillierte Auflistung der 161 im Projekt identifizierten unerwünschten Ereignisse wird zur weiteren Verwendung an die Tirol Kliniken übergeben werden.

Tabelle 5 zeigt die Verteilung der unerwünschten Ereignisse nach Schadensklassifikation und den Anteil der für das Indikatoren-Outcome kausalen Ereignisse für die im Projekt eingeschlossenen **Verdachtsfälle** gegenüber der Verteilung unter den **Kontrollfällen**.

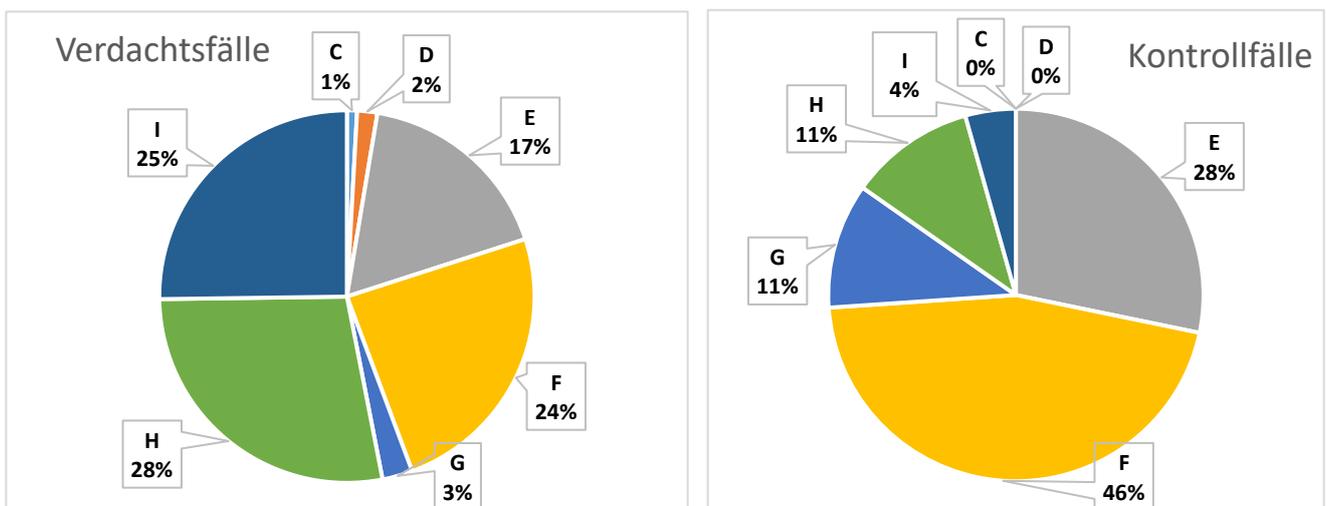
Tab. 5: Schadensklassifikation nach Verdachts- und Kontrollgruppe

Schadensklassifikationen	Verdachtsfälle (Indikator-positiv)		Kontrollfälle (Indikator-negativ)	
	Anzahl (%)	davon kausal	Anzahl (%)	davon kausal
C Ein Fehler, der den Patienten zwar erreicht, aber ihm keinen Schaden zugefügt hat.	0.87% (n=1)	0 -	0 -	- -
D Ein Fehler, der den Patienten erreicht hat und der eine Überwachung oder Intervention erforderte, um sicherzustellen, dass der Fehler dem Patienten keinen Schaden zugefügt hat.	1.74% (n=2)	0 -	0 -	- -
E Temporärer Patientenschaden, der eine Intervention erfordert.	17.39% (n=20)	40% (n=8)	28,26% (n=13)	7,69% (n=1)
F Temporärer Patientenschaden, der einen Krankenhausaufenthalt oder eine Verlängerung des Krankenhausaufenthaltes erfordert.	24.35% (n=28)	21,43% (n=6)	45,65% (n=21)	14,29% (n=21)
G Permanenter Patientenschaden.	2.61% (n=3)	66,67% (n=2)	10,87% (n=5)	20% (n=1)
H Patientenschaden, der eine lebenserhaltende Intervention notwendig macht.	27.83% (n=32)	87,50% (n=28)	10,87% (n=5)	20% (n=1)
I Patientenschaden, der zum Tod des Patienten beigetragen hat.	25.22% (n=29)	100% (n=29)	4,35% (n=2)	100% (n=2)
Σ	100% (n=115)	63,48% (n=73)	100% (n=46)	17,39% (n=81)

115 der identifizierten UE wurden in der Verdachtsgruppe, 46 in der Kontrollgruppe festgestellt. Bei der Gegenüberstellung der Schadensklassifikationen zeigt sich, dass in der Verdachtsgruppe die beiden schwersten Schadensklassifikationen H und I mehr als die Hälfte (53,05%) der UE ausmachen, während in der Kontrollgruppe der Anteil dieser beiden Klassen nur 15,22% ausmacht. Fast die Hälfte (45,65%) der UE in der Kontrollgruppe wurden mit F klassifiziert, haben also den Krankenhausaufenthalt für die PatientInnen, „nur“ verlängert. Abb. 3 zeigt die prozentuale Verteilung der Schadensklassifikationen in der Verdachts- und der Kontrollgruppe.

Während die Schäden der Klassifikation I sowohl in der Verdachts- als auch in der Kontrollgruppe zu 100% kausal für das Indikatoren-Outcome waren und damit zum Tod der PatientInnen, beigetragen haben, unterschied sich der prozentuale Anteil der kausalen UE für die anderen Klassifikationen zwischen den beiden Gruppen. Für die Verdachtsgruppe waren 87,50% (28) der UE der Klasse H und 66,76% (29) der Klasse G kausal, in der Kontrollgruppe waren es für beide Klassen jedoch nur 20% (1).

Abb. 3: Prozentuale Verteilung der Schadensklassifikationen



Diskussion

Zeil dieses Projektes war die Validierung der A-IQI anhand von Chart-Reviews mit dem GTT. Von den ursprünglich von den Tirol Kliniken übermittelten neunzehn auffälligen A-IQI aus dem Jahre 2019, wurden achtzehn Indikatoren (12 Indikatoren mit Erwartungswert bzw. österreichweitem Durchschnittswert als Zielbereich und 6 Sentinel-Indikatoren) als sicherheitsrelevant eingestuft und in die Analyse aufgenommen. Diese achtzehn Indikatoren (sechs konservative und zwölf chirurgische Indikatoren) umfassten insgesamt 4109 den Einschlusskriterien entsprechende Behandlungsfälle, von denen 252 als Stammdaten-positive Verdachtsfälle zur Auffälligkeit der Indikatoren beigetragen hatten. Entsprechend der oben beschriebenen Fallzahlenplanung wurden insgesamt 260 Kontrollen als Stichprobe aus dem entsprechenden Pool der Behandlungsfälle mit der gleichen Krankenhaushauptdiagnose gezogen. 91 Fallduplikate mussten aus der Analyse über alle Indikatoren (84 Duplikate in der konservativen und 7 in der chirurgischen Indikatorengruppe) entfernt werden. Ein Fall aus Indikator 01.10 musste während der Reviews aufgrund einer unzureichenden Dokumentation entfernt werden. Damit wurden insgesamt 421 Chart-Reviews, 233 in der konservativen, 188 in der chirurgischen Indikatorengruppe in die Auswertung aufgenommen. Da auf Ebene der einzelnen Indikatoren die Falldubletten inkludiert werden konnten, beträgt das kumulative n= 512.

Die prognostische Validität der auffälligen A-IQI wurde anhand einer Überprüfung der vorliegenden Patientendokumentationen auf vorliegende unerwünschte Ereignisse mit dem GTT bestimmt. Sensitivität, Spezifität sowie PPW und NPW der Indikatoren zeigten tendenziell gute bis moderate, in einzelnen Fällen perfekte, in wenigen Fällen mangelhafte Werte. Die Auswertung über alle Indikatoren scheint aufgrund der Größe der untersuchten Stichprobe (n= 421) eine gewisse Aussagekraft für alle Indikatoren zu haben. Die Sensitivität beträgt hier annähernd 70% (68,6%), während diese bei den einzelnen Indikatoren zwischen 50% und 100% liegt. Auch die Spezifität aller Indikatoren zusammen liegt bei knapp 70% (69,55%), bei den einzelnen Indikatoren zwischen 30,77% und 100%.

Der PPW liegt bei 36,65% über alle Indikatoren und annähernd 50% (47,62%) für die chirurgische Indikatorengruppe. Unter den konservativen Indikatoren finden sich vier (01.13, 01.43, 02.10 & 02.13) deren PPW unter 20% liegt. Im Gegensatz erreichten die A-IQI anhand der Überprüfung mit dem GTT sehr gute Ergebnisse bzgl. des NPW. Über alle Indikatoren waren fast 90% (89,62%) der A-IQI negativen Kontrollfälle tatsächlich ohne UE. Der zweitniedrigste NPW lag bei 80% (Indikator 01.13) während den mit Abstand niedrigsten NPW (57,14%) der Indikator 13.12 aufwies. Dieser Indikator, im stationären Aufenthalt erworbene Pneumonie, zeigte mit 75% wiederum den mit Abstand höchsten PPW unter den konservativen Indikatoren. Da PPW und NPW abhängig von der Prävalenz sind, lässt sich das Ausreißen dieses Indikators mit der veränderten Häufigkeit von bei dieser Hauptdiagnose vorliegenden UE erklären. Die für alle anderen

Indikatoren vermutete Prävalenz von ca. 10% wurde sicherlich überschritten, da per Definition das Auftreten eines Behandlungsfalles in diesem Indikator höchst wahrscheinlich mindestens ein UE, die stationär erworbenen Pneumonie, vorliegt.

Eine weitere Besonderheit bilden die vier „perfekten“ chirurgischen Indikatoren (15.34, 19.30, 37.10 & 43.30) mit einer Sensitivität, Spezifität sowie PPW und NPW von 100%. Da für die Indikatoren jeweils nur ein A-IQI positiver Fall vorlag, konnte die statistische Sättigung mit $n=5$ für die einzelnen Sample wohl nicht erreicht werden. Diese „perfekten“ Werte weisen auf eine insgesamt gute prädiktive Validität der vier Indikatoren hin, würden bei einer größeren Stichprobe voraussichtlich jedoch sinken. Bei der Überprüfung der $n=1$ Verdachts- und $n=4$ Kontrollfälle für den Indikator 40.11 wurde wiederum gar kein UE gefunden, weshalb die Aussagekraft der deskriptiven Auswertung dieses Indikators als gering anzusehen ist.

Für die Validität der A-IQI Indikatoren spricht weiterhin, dass der Anteil der Behandlungsfälle mit mindestens einem unerwünschten Ereignis mit zwei Ausnahmen (01.13 & 02.10), in der Indikator-positiven Verdachtsgruppe stets deutlich höher als in der Indikator-negativen Kontrollgruppe war. Mit Ausnahme des Indikators 13.12, wurde in den Kontrollgruppen, die aus Vorstudien vermutete Prävalenz von mindestens einem UE in 10% der Behandlungsfällen näherungsweise erreicht. In der Verdachtsgruppe wiederum fand sich im Durchschnitt über alle Indikatoren in mehr als einem Drittel (36,64%) der Behandlungsfälle mindestens ein unerwünschtes, in der chirurgischen Indikatorengruppe waren es annähernd 50% (47,62%).

Die verhältnismäßig breit gestreuten Ergebnisse bezüglich der entdeckten Behandlungsschäden pro 1000 Behandlungstage bzw. 100 stationäre Aufnahmen sind zwar vergleichbar mit Ergebnissen aus ähnlichen Studien (Becker 2015), sie erlauben jedoch keine verbindliche Aussage zur Validität der A-IQI. Jedoch können diese Daten als Grundlage für einen Vergleich zur Entwicklung der Indikatoren innerhalb der Tirol Kliniken über mehrere Jahre dienen.

Bei der Auswertung der identifizierten Patientenschäden zeigt sich in der Fallgruppe neben der höheren Proportion von Behandlungsfällen mit mindestens einem UE auch ein häufigeres Auftreten von schweren Schadensklassifikationen. Während in der höchsten Schadensklasse I sowohl in der Verdachts- als auch in der Kontrollgruppe 100% der UE kausal für das Indikatoren-Outcome waren, ist der prozentuale Anteil an kausalen UE in allen anderen Schadensklassen in der Verdachtshöhe deutlich höher.

Die Art der detektierten Schäden entspricht im Wesentlichen den in der Literatur beschriebenen Ergebnissen. Annähernd ein Drittel (32,30%) der unerwünschten Ereignisse waren im Krankenhaus erworbene Infektionen, von denen wiederum fast ein Drittel (32,69%) Pneumonien ausmachten. Knapp 20% (19,88%) waren unterschiedliche Schäden wie Beinvenenthrombosen und/oder Pulmonalembolien, ein intra- oder postoperativer Myokardinfarkt oder ein postoperatives Lungenödem. Weiter häufige

Schadenskategorien waren chirurgische Fehler, Pflegefehler und (Nach)Blutungen, während nur wenige (3,73%) medikamentöse Behandlungsfehler entdeckt wurden. Die Kategorisierung der identifizierten unerwünschten Ereignisse kann als Grundlage für detailliertere Analysen im Rahmen des innerklinischen Qualitätsmanagements dienen.

Limitationen

Wie alle Prädiktionsstudien unterliegt diese Studie verschiedenen Limitationen. In Bezug auf die Fallzahlschätzungen und Auswahl von Prädiktoren bezieht die vorliegende Studie ihre Annahmen aus vergangenen Daten (hier aus dem Jahr 2019). Sollte sich das Geschehen in den Folgejahren grundsätzlich ändern, so muss damit gerechnet werden, dass die in dieser Studie untersuchten Prädiktoren eventuell nicht signifikant zur Prädiktion beitragen werden.

Insgesamt sind im A-IQI Set des BMSGPK 2019 380 Qualitätsindikatoren enthalten, davon 145 Indikatoren mit Zielbereich und 17 Sentinel-Indikatoren. Auf Grund der Systematik des A-IQI Indikatorensystems gelten die einzelnen Indikatoren immer nur für definierte Behandlungsfälle, z.B. für eine bestimmte Diagnose oder medizinische Leistung und für eine bestimmte Altersgruppe. Im Gegensatz zu allgemeinen Indikatoren, wie den Indikatoren der Agency for Healthcare Research an Quality (AHRQ), können spezifische Indikatoren daher immer nur für jene Fälle ausgewertet werden, für die sie gelten. In Innsbruck waren nur ca. 10 % der Indikatoren mit Zielbereich bzw. Sentinel-Indikatoren positiv und konnten somit in die Analyse einbezogen werden. Über die Validität der im Jahr 2019 unauffälligen Indikatoren in Bezug auf die Prädiktion von unerwünschten Ereignissen kann daher keine Aussage getroffen werden. Eine umfassende Validierung des gesamten A-IQI Systems würde jedoch eine im Rahmen einer monozentrischen Studie nicht praktikable Anzahl von systematischen Analysen von Krankengeschichten erfordern.

Für die Indikatoren 15.34, 19.30, 37.10, 40.11 und 43.30 war nur jeweils ein positiver Fall (rote Ampel) für das Jahr 2019 vorhanden. Davon zeichneten sich die Indikatoren 15.34 und 40.11 durch eine Stammdatengesamtzahl von nur $n=21$ bzw. $n=16$ aus. Dies sind also seltene Indikatoren. Dies kann dazu führen, dass der jeweilige Indikator in Bezug auf das absolute Ausmaß in der Zukunft eine nur geringen bis gar keine Bedeutung hat. Andererseits sollten diese Indikatoren gerade deshalb mit einbezogen werden, da die relative Häufigkeit für Positivität mit 5% jedenfalls nicht zu vernachlässigen ist. Zudem sind vier dieser Indikatoren (15.34, 19.30, 37.10 und 43.30) Sentinel-Indikatoren, bei denen definitionsgemäß mit hoher Wahrscheinlichkeit von einem kausalen unerwünschten Ereignis ausgegangen wird.

Eine weitere Einschränkung ergibt sich aus der Tatsache, dass das Risiko eines unerwünschten Ereignisses von vielen Covariaten abhängt (wie z.B. Alter, Komorbiditäten, sowie aggressive bzw. konservative Indikationsstellung, Patientenverfügung, OP nur an diesem KH angeboten u.a.). Die A-IQI Indikatoren sind

nur nach Alter und Geschlecht standardisiert. Da nicht sicher ist, ob andere Covariate bei indikatorpositiven Fällen und Kontrollen gleich verteilt sind, ist die kausale Interpretation der Ergebnisse wie bei allen Prädiktionsmodellen eingeschränkt.

Eine weitere Limitation könnte in der Art der Durchführung der Chart-Reviews in diesem Projekt begründet sein. Zunächst haben alle am Projekt beteiligten Reviewer erstmals mit dem GTT gearbeitet, und in der Review Phase zeigte sich, dass mit zunehmender Erfahrung die Reviewer sicherer und zielgenauer mit diesem Tool umzugehen wussten. Darüber hinaus wurden die Reviews ausschließlich im elektronischen Dokumentationssystem der Tirol Kliniken und nicht anhand papierbasierter Patientenakten durchgeführt. Da es in der elektronischen Akte zusätzlicher Zeit zur Orientierung im Behandlungsfall sowie zum Öffnen der digital hinterlegten Dokumente bedarf, konnten die durch die GTT-Methodik vorgeschriebenen 20 Minuten pro Chart-Review nicht immer eingehalten werden. Dies könnte zum einen dazu geführt haben, dass die Akten genauer als vorgesehen durchgeschaut und dadurch mehr UE als üblich festgestellt wurden. Andererseits konnten einzelne Dokumente, wie die eingescannten Medikationsblätter, aufgrund eingeschränkter Bildqualität nur unzureichend gescreent worden sein, was zu einer reduzierten Detektionsrate von medikamentösen Behandlungsfehlern geführt haben könnte.

Schließlich war es nicht möglich, die Reviewer und Supervisoren während ihrer Arbeit an den Chart-Reviews vollständig zu verblinden. Bei sechzehn der achtzehn inkludierten Indikatoren war der Outcome „Anteil Todesfälle“. So hätten die Reviewer und Supervisor während der Chart-Reviews mutmaßen können, dass die PatientInnen, die den entsprechenden Krankenhausaufenthalt überlebt hatten mit einer hohen Wahrscheinlichkeit zur Kontrollgruppe gehörten. Diese unvollständige Verblindung konnte aufgrund der für die Chart-Reviews zwingend notwendigen Akteneinsicht nicht verhindert werden und hat möglicherweise die Beurteilung in einzelnen Fällen beeinflusst.

Schlussfolgerungen

Die oben dargestellten Ergebnisse zeigen, dass durch die systematische Suche nach unerwünschten Ereignissen durch standardisierte zweistufige Analyse von Krankengeschichten eine mit der Detektionsrate in früheren Studien vergleichbare Anzahl von Behandlungsschäden gefunden wurde.

Behandlungsfälle mit auffälligem A-IQI Indikator wiesen signifikant mehr unerwünschte Ereignisse und ebenso signifikant mehr schwere und für das Outcome kausale unerwünschte Ereignisse auf als Kontrollfälle ohne auffälligen Indikator. Die getesteten A-IQI scheinen damit eine gute prädiktive Validität für die Identifikation von klinisch relevanten behandlungsbedingten unerwünschten Ereignissen aufzuweisen.

Statistisch signifikant waren die Unterschiede in der Verteilung der unerwünschten Ereignisse für den Test

über alle Indikatoren, die zwei untersuchten Indikatorengruppen (chirurgische und konservative Indikatoren), sowie für sechs einzelne Indikatoren.

Im Rahmen des klinischen Qualitäts- und Risikomanagements im Krankenhaus wird angestrebt, möglichst viele okkulte unerwünschte Ereignisse zu detektieren, um entsprechende Präventionsmaßnahmen ableiten zu können. Die Autoren dieses Berichtes empfehlen daher, für die Auswahl von Behandlungsfällen, die einer systematischen Analyse der Krankengeschichten mit dem Global Trigger Tool oder ähnlichen Verfahren unterzogen werden, Fälle mit positivem A-IQI Indikator statt Zufallsstichproben zu verwenden, da durch dieses Vorscreening mehr sowie häufiger schwere und für das Outcome kausale behandlungsbedingte Patientenschäden detektiert werden können.

Die weitere Forschung in diesem Bereich sollte sich darauf konzentrieren, eine möglichst multizentrische Validierung der A-IQI zu unternehmen, um eine größere Anzahl von Indikatoren zu testen sowie mögliche Rückschlüsse auf die Anwendbarkeit der A-IQI für einen überinstitutionellen Vergleich von Krankenhäusern zu erlauben.

Referenzen

- Anderson O, Davis R, Hanna GB, Vincent CA: (2013) Surgical adverse events: A systematic Review. *AM J. Surgery*. 2013 Aug; 206:253-62
- Becker A (2015): Das Global Trigger Tool des Institute for Health Care Improvement zum Screening auf Patientenschäden. In A. Becker (Hrsg.), *Beiträge zur Patientensicherheit im Krankenhaus*. (S. 47-64). KU-Gesundheitsmanagement
- Brennan TA, Leape LL, Laird NM, Hebert I, Localio AR, Lawthers AG, Newhouse JP, Weiler PC, Hiatt HH: Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study. *N Engl J Med* 1991; 324: 370-376.
- Bundesministerium für Soziales, Gesundheit, Pflege und Konsumentenschutz (2019): Austrian Inpatient Quality Indicators (A-IQI). Bericht 2019.
<https://www.sozialministerium.at/Themen/Gesundheit/Gesundheitssystem/Gesundheitssystem-und-Qualitaets-sicherung/Ergebnisqualitaetsmessung.html>. Zugriff am 2.12.2020.
- Chow S, Shao J, Wang H: *Sample size calculations in clinical research*, second edition. Chapman & Hall/CRC. Boca Raton, Florida (2008).
- D'Agostino R, Chase W, Belanger A: The appropriateness of some common procedures for testing the equality of two independent binomial populations'. *Am Stat* 1988;42(3):198-202.
- Farquhar M: AHRQ Quality Indicators. Chapter 45: In: *Patient Safety and Quality: An Evidence-Based Handbook for Nurses*. Vol. 3, 41–67.
- Fleiss JL, Levin B, Paik MC: *Statistical Methods for rates and proportions* (third edition). John Wiley & Sons. New York (2003).
- Griffin FA, Resar RK: IHI Global Trigger Tool for measuring adverse events (second edition). IHI Innovation Series white Paper. Institute for Healthcare Improvement. Cambridge, Massachusetts (2009).
- Health Quality & Safety Commission New Zealand: *The Global Trigger Tool. A practical implementation guide for New Zealand District Health Boards* (2012).
- Herold A: Das Global Trigger Tool. Ein Messinstrument der Patientensicherheit. In: *Passion Chirurgie* 2013;3(05). Online verfügbar unter <https://www.bdc.de/safety-clip-das-global-trigger-tool-messinstrument-der-patientensicherheit/>.
- Hoffmann-Völkig, Kästenbauer T, Mück U, Zottl M, Huf W, Ettl B: Detektion unerwünschter Ereignisse mittels IHI Global Trigger Tool im Kontext der Einführung eines Risikomanagement-Systems: Eine retrospektive Studie über einen Zeitraum von 3 Jahren an der Herz- und Gefäßchirurgischen Abteilung eines Wiener Schwerpunktkrankenhauses. *Z Evid Fortbild Qual Gesundheitswes* 2018;131:38-45.
- Lachin JM: *Biostatistical Methods*. John Wiley & Sons. New York (2000).
- Machin D, Campbell M, Fayers P, Pinol A: *Sample Size Tables for Clinical Studies* (2nd edition). Blackwell Science. Malden, Mass. (1997).
- Makary MA, Daniel M: Medical error- the third leading cause of death in the US. In: *BMJ* 2016; 353:i2139. DOI: 10.1136/bmj.i2139.

Murff HJ, Patel VL, Hripcsak G, Bates DW: Detecting adverse events for patient safety research: a review of current methodologies. *Biomed Inform* 2003;36(1-2):131–143. DOI: 10.1016/j.jbi.2003.08.003.

Ryan TP: *Sample size determination and power*. John Wiley & Sons. Hoboken, New Jersey (2013).

Schrapppe M: *APS-Weißbuch Patientensicherheit. Sicherheit in der Gesundheitsversorgung: Neu denken, gezielt verbessern*. Berlin: Medizinisch Wissenschaftliche Verlagsgesellschaft (2018).

De Vries EN, Ramrattan MA, Smorenburg SM, Gouma DJ, Boermeester MA: The incidence and nature of in-hospital adverse events: a systematic review. *Qual & Saf Health Care* 2008; 17(3):216–223. DOI: 10.1136/qshc.2007.023622.

Zúñiga F, Blatter C; Wicki R, Simon M: Nationale Qualitätsindikatoren in Schweizer Pflegeheimen: Fragebogenerhebung zur Zuverlässigkeit der Daten und Nützlichkeit aus Sicht der Anwender. *Z Gerontol Geriatr* 2019;52(8):730–736. DOI: 10.1007/s00391-019-01583-7.

Anhang

- A) Case Report Form
- B) Fallzahlberechnungen und Entfernung der Duplikate per Indikator

Anhang A – Case Report Form

ID:					
Modul	Klassifizierung des Schadens (E-I)	Beschreibung des Vorfalls/Anmerkungen	Modul	Klassifizierung des Schadens (E-I)	Beschreibung des Vorfalls/Anmerkungen
		Trigger - Chirurgie			
		Trigger - Generelle Behandlung			
G1		Bluttransfusion oder Transfusion von	K1		Erneute Operation
G2		Ruf des Herzalarm-Teams oder Herz- und Atemstillstand und starke Verschlechterung des Patienten	K2		Änderung des Eingriffes
G3		Akute Dialyse	K3		Verlegung auf die Intensivstation nach der Operation
G4		Positive Blutkulturen	K4		Intubation/Reintubation/BI PAP/CPAP im Aufwachraum
G5		Röntgen- oder Ultraschalluntersuchungen bzgl. Embolie oder Phlebothrombose	K5		Röntgenuntersuchung während der Operation oder im Aufwachraum
G6		Plötzliches Fallen des Hb oder Hkt \geq 25%	K6		Tod während oder direkt nach der Operation
G7		Sturz des Patienten	K7		Postoperative Beatmung > 24 Stunden
G8		Dekubitus	K8		Intra-operative Gabe von Epinephrin, Norepinephrin, Naloxon oder Flumazenil
G9		Wiederaufnahme innerhalb von 30 Tagen	K9		Post-operativer Anstieg des Troponin Niveaus über Normwert
G10		Fixierung des Patienten	K10		Änderung der Form der Anästhesie während der Operation
G11		Im Krankenhaus erworbene Infektionen	K11		Überwachung durch einen anderen Spezialisten im Aufwachraum
G12		Schlaganfall im Krankenhaus	K12		Falsche pathologische Befunde, die nicht in Verbindung zur prä-operativen Diagnose stehen
G13		Verlegung auf ein anderes Pflege- oder Behandlungsniveau	K13		Anlegen eines Arterienkatheters oder eines ZVK während der Operation
G14		Komplikationen	K14		Operationszeit > 8 Stunden
G15		Sonstiges	K15		Entfernung, Schaden oder Reparatur von Schäden an Organen während der Operation
		Trigger - Medikamente			
M1		Clostridium difficile in der Stuhlprobe	Trigger - Intensivbehandlung		
M2		PTT > 100 Sekunden	I1		Pneumonie
M3		INR > 4	I2	1	Wiederaufnahme auf der Intensivstation
M4		Glukose < 3,0 mmol/l bzw. < 54mg/dl	I3		Behandlung/Eingriff auf der Intensivstation
M5		Anstieg des Harnstoff/Creatinin, 2fach über Normalwert	I4		Intubation/Re-Intubation
		Trigger - Perinatal			
M6		Vitamin K	P1 a)		Apgar score < 7,0 nach 5 Minuten
M7		Antihistaminika	P1 b)		Nabelschnur pH < 7,05 und/oder BE > 10
M8		Flumazenil (Anexate)	P2		Transport oder Verlegung von Mutter oder Kind
M9		Naloxon	P3		Magnesiumsulfat oder Ebricanyl
M10		Antiemetika	P4		Dammverletzungen 3. oder 4. Grades
M11		Hypotonie/Übersiedlerung	P5		Einleitung der Geburt
M12		Plötzlicher Stopp der Medikation	P6		B-Glukose < 2,2mmol/l bzw. < 40mg/dl beim Kind
		Trigger - Notaufnahme			
		Anzahl Trigger			
		Anzahl Patientenschäden			
		Zeitraum			
		Anzahl Behandlungstage			
				A1 Wiederaufnahme in der Notaufnahme innerhalb von	
				A2 Aufenthalt in der Notaufnahme > 6 Stunden	

Anhang B – Fallzahlberechnungen und Entfernung der Duplikate per Indikator

Gruppe		Stamdate n gesamt	Indikator- positiv (V)	Fallzahl Kontrollen (K)	Fälle in Chart	Duplikate in V	Chart Reviews	
01 - Herzinfarkt								
01.10	Hauptdiagnose Herzinfarkt (Alter > 19), Anteil Todesfälle	819	68*	35	103	34	69	Konservative Indikatoren (n=233)
01.13	Hauptdiagnose Herzinfarkt, Altersgruppe 65-84, Anteil Todesfälle	436	30	15	45	6	39	
01.43	Hauptdiagnose transmuraler Herzinfarkt (Alter > 19), Anteil Todesfälle	455	41	21	62	32	30	
02 - Herzinsuffizienz								
02.10	Hauptdiagnose Herzinsuffizienz (Alter > 19), Anteil Todesfälle	188	24	12	36	9	27	
02.13	Hauptdiagnose Herzinsuffizienz, Altersgruppe 65-84, Anteil Todesfälle	105	12	17	29	2	27	
13 - Lungenentzündung (Pneumonie)								
13.12	Pneumonie im stat. Aufenthalt erworben (hospital-acquired pneumonia), Anteil Todesfälle	105	28	14	42	1	41	
15 - Operationen an der Lunge (große thoraxchirurgische Eingriffe)								
15.34	Teilresektionen der Lunge (ohne Lobektomie) bei Bronchialkarzinom, Anteil Todesfälle	21	1	4	5	0	5	Chirurgische Indikatoren (n=188)
16 - Laparoskopische Cholezystektomien								
16.60	Laparoskopische Cholezystektomien ohne Tumorfälle, Anteil Umsteiger	239	12	17	29	0	29	
16.70	Laparoskopische Cholezystektomien ohne Tumorfälle, Anteil ERCP nach OP-Tag	227	7	28	35	0	35	
19 - Magenoperationen								
19.30	Adipositas-Chirurgie, Anteil Todesfälle	57	1	4	5	0	5	
36 - Wechsel einer Hüftgelenkendoprothese								
36.10	Hüft-Endoprothesen-Revisionen insgesamt, Anteil Todesfälle	208	5	20	25	0	25	
37 - Kniegelenkersatz								
37.10	Knie-Endoprothesen-Erstimpl. b. Gonarthrose u. chronischer Arthritis, Anteil Todesfälle	302	1	4	5	0	5	
38 - Wechsel einer Kniegelenkendoprothese								
38.10	Knie-Endoprothesen-Revisionen insgesamt, Anteil Todesfälle	73	2	8	10	0	10	
40 - Hüftgelenknahe Frakturen								
40.11	Hüftgelenknahe Fraktur, Altersgruppe 20-59, Anteil Todesfälle	16	1	4	5	0	5	
40.40	Pertrochantäre Fraktur (Alter > 19), Anteil Todesfälle	130	10	21	31	5	26	

40.46	Pertrochantäre Fraktur, Altersgruppe >= 90, Anteil Todesfälle	36	6	24	30	2	28
<i>41 - Operationen der Wirbelsäule und des Rückenmarks</i>							
41.12	Operationen an der Wirbelsäule und am RM (ohne Tumor, Trauma), Anteil Todesfälle	1433	2	8	10	0	10
<i>43 - Entfernung der Niere (Nephrektomie und partielle Nephrektomie)</i>							
43.30	Partielle Nephrektomien bei bösartiger Neubildung der Niere, Anteil Todesfälle	78	1	4	5	0	5
Summe		4109	252	260	512	91	421

* ein Fall aufgrund zu geringer Datenlage (fehlende Dokumentation) ausgeschlossen